



UNIVERSITÉ D'ARTOIS

Avis de Soutenance

Monsieur Alexis DE COLNET

Informatique et applications

Soutiendra publiquement ses travaux de thèse intitulés

Des fonctions difficiles en compilation de connaissances : bornes inférieures et applications

dirigés par Monsieur Pierre MARQUIS

Soutenance prévue le **lundi 26 septembre 2022** à 9h00

Lieu : UFR des Sciences Jean Perrin Rue Jean Souvraz SP 18 F-62307 Lens Cedex France

Salle : des thèses

Composition du jury proposé

M. Pierre MARQUIS	Université d'Artois	Directeur de thèse
M. Stefan SZEIDER	Technische Universität Wien	Rapporteur
M. Florent CAPELLI	Université de Lille	Examinateur
M. Adnan DARWICHE	Universtiy of California	Rapporteur
Mme Sophie TISON	Université de Lille	Examinatrice
M. Stefan MENGEL	CNRS - Université d'Artois	Co-directeur de thèse

Résumé :

Le thème de la thèse est la compilation de connaissances, une approche pour la résolution de problèmes difficiles à résoudre du point de vue du calcul et qui vise à réduire cette complexité en pré-traitant (en compilant) une partie du problème connue à l'avance et modélisée par une fonction. Il s'agit de trouver une représentation de la fonction dans une classe de représentations appelée langage de compilation. Pour un langage L donné, la taille d'une fonction dans L désigne la taille de sa plus petite représentation dans L . Dans cette thèse, nous étudions différents aspects de la compilation dans le langage DNNF, le langage où les fonctions Booléennes sont représentées par des circuits sous forme normale négative décomposable (DNNF). Dans la première partie de la thèse, nous montrons des bornes inférieures sur la taille de fonctions particulières dans le langage DNNF. Pour certaines fonctions, notamment des contraintes pseudo-Booléennes, nous obtenons des bornes inférieures par l'application de techniques usuelles pour l'analyse de circuits en DNNF. Pour d'autres fonctions, ces mêmes techniques s'avèrent insuffisantes. En particulier, pour les formules de Tseitin, qui sont des formules CNF représentant des systèmes d'équations linéaires structurés par des graphes, nous avons amélioré les techniques existantes pour montrer une borne exponentielle sur leur taille dans le langage DNNF. Quand la taille d'une fonction dans un langage est trop grande pour envisager sa compilation, on peut tenter de compiler une approximation de la fonction, on parle alors de compilation de connaissances approchée. Nous avons étudié deux notions d'approximation qui offrent des garanties sur l'erreur d'approximation. Pour chacune de ces notions, nous avons trouvé des exemples de fonctions dont toutes les approximations ont une taille trop grande dans de nombreux langages de compilation. Dans la seconde partie de la thèse, nous donnons des applications des nos bornes inférieures. Ces applications font le lien entre la compilation de connaissances et le domaine de la complexité des preuves. Notamment, nous utilisons notre borne inférieure sur la taille des formules de Tseitin dans le langage DNNF pour obtenir une caractérisation des formules de Tseitin non-satisfiables (dont les graphes sont de degré borné par une constante) pour lesquelles il existe des réfutations par résolution dites régulières de taille polynomiale en le nombre de variables. Une seconde application concerne à la fois certains systèmes de preuve et les compilateurs utilisant la compilation dite ascendante. La compilation ascendante a pour particularité qu'elle construit le circuit représentant la fonction initiale en combinant des circuits intermédiaires. Dans le cas de formules non-satisfiables, la séquence de circuits intermédiaires générés lors d'une compilation ascendante peut être vue comme une réfutation de la formule. Nous analysons des cas de compilations ascendantes de formules (pour certaines non-satisfiables) ayant de petites représentations dans le langage cible, mais pour lesquelles des circuits intermédiaires de taille exponentielle sont inévitables. Dans les derniers chapitres de la thèse, nous tentons d'élargir le champ de recherche en compilation de connaissances tout en restant dans l'esprit de la « carte de compilation de connaissances », un modèle proposé par Darwiche et Marquis pour comparer les langages de compilation en termes de compacité (quels sont les langages permettant d'obtenir les plus petites représentations ?) et en termes d'efficacité calculatoire (pour quels langages existe-t-il des algorithmes en temps polynomial pour répondre aux requêtes ?). Nous introduisons de nouvelles requêtes pour les langages Booléens, plus précisément des requêtes d'énumération, et nous initions des « cartes de compacité » pour des langages pour des fonctions non-Booléennes.